Prediction of Neonatal Jaundice using Fuzzy Clustering Methods

Hadi Talebian Moghaddam*, Farzad Towhidkhah, Rasoul Khayati, Abdulhamid Ghlichkhani, Mohammad Torkaman, Zohreh Kavehmanesh

Abstract— **Jaundice is the most common and one of the most vexing problems that can occur in the newborn. Although most jaundiced infants are otherwise perfectly healthy, they make us anxious because bilirubin is potentially toxic to the central nervous system and, if the serum bilirubin level is very high, kernicterus (bilirubin encephalopathy) can occur.**

The idea of predicting risk of neonatal jaundice before and after delivery is tried in this paper. For this purpose some different fuzzy clustering methods including fuzzy c-means (FCM), subtractive clustering (SC) and fuzzy adaptive resonance theory mapping (fuzzy ARTMAP) are used. A total of 552 medical records were collected from newborns in two general hospitals in Tehran, Iran in 2006. To evaluate results of the applied methods we used performance evaluation matrix criteria, which include percentage of accuracy (correct classification) (CC%), sensitivity (SE%), and specifity (SP%). The above mentioned criteria for jaundice prediction before delivery were approximately 76%, 97%, 56%, **while for jaundice prediction after delivery were** 81%, 88%, 67%, **respectively. These results show that the proposed systems can achieve satisfying results for predicting risk of jaundice considering this fact that physicians do not have any estimation about probability of jaundice appearance.**

Keywords: Fuzzy Clustering, Neonatal Jaundice, Prediction, Simplified fuzzy ARTMAP (SFAM), Fuzzy C-Means (FCM), Subtractive Clustering

I. Introduction

Jaundice is a common and, in most cases, benign problem in neonates. In 1990s jaundice was the prevalent cause of returning of newborns to hospitals [1]. So, early prediction of newborns liable to jaundice is important.

Hemolytic disease of the newborn is a common cause of neonatal jaundice. Nonetheless, because of the immaturity of the pathways of bilirubin metabolism, many newborns suffer from jaundice without evidence of hemolysis.

Untreated, severe indirect hyperbilirubinemia is potentially neurotoxic. Jaundice is observed during the first week of life in approximately 60% of term infants and 80% of preterm infants. The color usually results from the accumulation of bilirubin pigment under skin. It may also be somewhat due to deposition of the pigment after it has been converted to conjugated bilirubin in the liver cell. The unconjugated form is neurotoxic in infants at certain concentrations and under various conditions. Conjugated bilirubin is not neurotoxic but indicates a potentially, serious disorder.

Compared with adults, newborn infants have a twofold to threefold greater rate of bilirubin production. This is caused in part by an increased red blood cell mass (higher hematocrit) and a shortened erythrocyte life span of 70-90 days, compared with the 120 day erythrocyte life span in adults.

Risk factors for indirect hyperbilirubinemia include maternal diabetes, race, prematurity, drugs, male sex, oxytocin induction, breast-feeding, and a sibling who had jaundice. A family history of neonatal jaundice, exclusive breast-feeding, Asian race, and maternal age older than 25 year identify approximately 60% of cases of extreme hyperbilirubinemia [2].

In what follows, in section II comes a review of previous works on jaundice. Then a summary of the applied clustering methods is given in section III. Section IV includes data gathering and feature selection matters. Evaluation method comes in section V. In section VI the results of the applied methods are shown. Discussion and conclusion come at last in sections VII and VIII, respectively.

II. Literature Study

Lee et al proposed a model of bilirubin circulation in the body using mathematical equations [3]. In another research in 2002, Seyfang et al presented an algorithm for diagnosis and treatment of jaundice in newborns using ASBRU language [4]. This algorithm presents how to diagnose and treat jaundice according to some laboratorial data including mother's and child's blood types.

American Academy of Pediatrics also suggested an algorithm for diagnosis and treatment of jaundice [5]. Another research in Switzerland shows the effect of season on the amounts of bilirubin in 5540 newborns cured by phototherapy [6]. According to this research, the maximum number of phototherapies occurred during May to August.

In 2001 Stevenson et al suggested the use of End-tidal carbon monoxide for ambient (ETCOc), to predict jaundice

Paper submitted September 1, 2006.

Hadi Talebian Moghaddam is with the department of bioelectrical engineering, biomedical engineering faculty, Amirkabir university of technology, P.O.Box 19395-5999 Tehran, Iran (corresponding author to provide phone: +98-21-8883-2990#2240; fax: +98-21-8881-2180; e-mail: htmoghaddam@yahoo.com).

Farzad Towhidkhah is as an assistant professor with the department of bioelectrical engineering, biomedical engineering faculty, Amirkabir university of technology, Tehran, Iran (e-mail: towhidkhah@aut.ac.ir).

Rasoul Khayati is with the department of bioelectrical engineering, engineering faculty, Shahed university, Tehran, Iran (e-mail: khayati@shahed.ac.ir).

Abdulhamid Ghlichkhani received the MSc degree in bioelectrical engineering from biomedical engineering faculty, Amirkabir university of technology, Tehran, Iran (e-mail: hghlich@gmail.com).

Mohammad Torkaman is as a Neonatalogist with the department of pediatrics, Baqiyatallah medical sciences university, Tehran, Iran (e-mail: m_d_torkaman@yahoo.com).

Zohreh Kavehmanesh. is as a Pediatric Gastroentrologist with the department of pediatrics, Baqiyatallah medical sciences university, Tehran, Iran (e-mail: z_kaveh@hotmail.com).

risk in the first days after the birth considering release of carbon monoxide when heme changes to bilirubin [7].

Khayati presented a procedure for jaundice treatment using Stateflow and Simulink toolboxes by defining different states, relations between them, effects of different inputs and determining related outputs [8].

In 2003 Ahouraei et al tried to diagnose neonatal jaundice using an MLP neural network [9].

III. CLUSTERING METHODS

The objective of cluster analysis is the classification of objects according to similarities among them, and organizing of data into groups. Clustering techniques are among the unsupervised methods, they do not use prior class identifiers.

A. Fuzzy C-means Algorithm

The fuzzy c-means clustering method is the most widely used algorithm to detect constrained fuzzy c-partitions. A constrained fuzzy c-partition can be briefly described as follows: Let $X = \{x_1, x_2, ..., x_n\}$ be a set of n unlabeled feature data vectors with $x_k \in \Re^p (1 \le k \le n)$, and c $(2 \le c \le n)$ be a number of fuzzy subsets (clusters) defined in X: Given that the membership function of the *k*th vector to the *i*th cluster is denoted as:

$$u_{ik} = \{u_i(x_k), 1 \le i \le c, 1 \le k \le n\}, \quad (1)$$

the c fuzzy clusters constitute a constrained fuzzy c-partition in X if the next three conditions are satisfied,

$$0 \le u_{ik} \le 1, \quad \forall i,k \quad 2(a)$$

$$0 \le \sum_{k=1}^{n} u_{ik} \le n, \quad \forall i \quad 2(b)$$

$$\sum_{k=1}^{n} u_{ik} = 1, \quad \forall k \quad 2(c)$$

Whenever the last condition, (2c), is not satisfied the fuzzy c-partition is said to be unconstrained. Fuzzy c-means is able to detect constrained fuzzy c-partitions by minimizing an objective function. This optimization procedure is done by minimizing Eq. (3), while keeping Eq. 2(c).

$$J_m(U,V;X) = \sum_{k=1}^{n}\sum_{i=1}^{c} (u_{ik})^m \|x_k - v_i\|_A^2 \quad (3)$$

where $U = \{[u_{ik}], 1 \le i \le c, 1 \le k \le n\}$ is the partition matrix, $V = \{[v_i], 1 \le i \le c\}$ with $v_i \in \Re^p$ is the vector of the resulted cluster centers, $m \in (1, \infty)$ is a factor to adjust the membership degree weighting effect, and $\|.\|_A$ is any inner product norm. The cluster centers and the respective membership functions that solve the constrained optimization problem in (3) are given by Eq. (4) and (5).

$$v_i = \frac{\sum_{k=1}^{n} u_{ik}^m x_k}{\sum_{k=1}^{n} u_{ik}^m}, \quad 1 \le i \le c \quad (4)$$

Eqs. (4) and (5) constitute an iterative optimization procedure [10].

$$u_{ik} = \frac{1}{\sum_{j=1}^{c} \left(\frac{\|x_k - v_i\|_A}{\|x_k - v_j\|_A} \right)^{2/m-1}}, \ 1 \le i \le c, 1 \le k \le n \quad (5)$$

B. Simplified Fuzzy ARTMAP System (SFAM)

The structure of the SFAM is shown in Fig. 1. The input vectors are first complement coded to become vectors ***I*** which are applied to the input layer as shown. Each node in the output category layer is linked through a set of top-down weights to each node in the input layer. The *N* nodes in the output category layer label the *M* categories or classes that the SFAM has to learn to recognize, usually *N* > *M*. When active during testing, an output category node indicates the class by pointing to the corresponding category classification node. The vigilance parameter ρ has to be chosen to determine the number of classes found. Match tracking causes automatic adjustment of ρ if classification errors are found in training. The activities are calculated as $T_j(\boldsymbol{I})$ where

$$T_j(I) = \frac{|I \wedge w_j|}{\alpha + |w_j|} \quad (6)$$

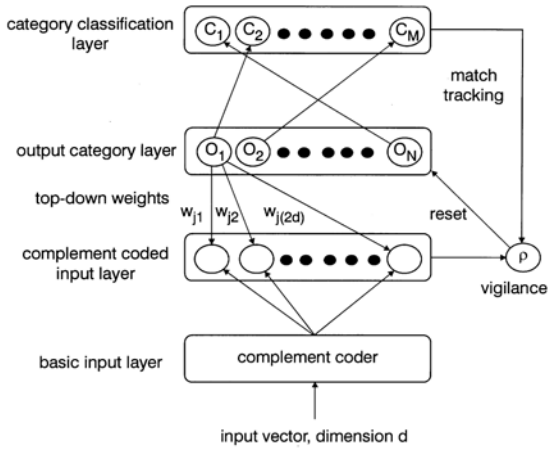

Fig. 1. Simplified fuzzy ARTMAP architecture [11]

Here $\wedge$ is the fuzzy AND operator. The activity $T_j(I)$ measures how well w_j represents the class of *I*. The node with the largest activation, known as the winning node, is designated with the suffix *J*. The value of α is fixed typically at 0.000001. The degree of match between the output category node and an input vector is given by the match function, $MF(I,W_J)$ defined as

$$MF(I, w_J) = \frac{|I \wedge w_J|}{|I|} = \frac{|I \wedge w_J|}{d} \quad (7)$$

This function measures to what extent ***I*** belongs to the class represented by $\boldsymbol{w}_J$. If $MF(I, W_J) \ge \rho$, and the winning node represents the same class as that to which ***I*** belongs (category agreement), then its weight vector w_J is updated according to

$$w_J^{new} = (I \wedge w_J^{old}) \quad (8)$$

If $MF(I, W_J) \ge \rho$, but the output category node *J* does not represent the class to which ***I*** belongs (category

disagreement), a different node must be used to represent ***I***, or eventually a new node may need to be assigned. The vigilance is increased to exceed the match function value by a small amount. Once a node has been assigned to the training vector, the vigilance is reset to its starting value. During testing each test vector is applied in turn and its class is predicted as that of the node which exhibits the highest activation [11].

C. Subtractive Clustering

In this paper, Chiu's subtractive clustering is applied [12]. Subtractive clustering is an efficient algorithm, which does not require optimization, being for this reason a good choice for the initialization of neuro-fuzzy networks. Chiu's algorithm belongs to the class of potential function methods which a set of points are defined as possible group centers, each of them being interpreted as an energy source [13].

Let Z^N be a set of N data samples, z_1, z_2,…, z_N. As it was referred, it is admitted that each of the samples defines a possible cluster center. The potential associated to z_i is (9):

$$P_i(z_i, Z^N) = \sum_{j=1}^{N} e^{-\alpha\|z_i - z_j\|^2}, \alpha = \frac{4}{r_a^2}, i = 1,2,...,N \quad (9)$$

where $r_a > 0$ is the radii parameter, a constant which defines the neighborhood radius of each point. Thus, points z_j located out of the radius of z_i will have a smaller influence in its potential. On the other hand, the effect of points close to z_i will grow with the proximity. In this way, points with a dense neighborhood will have higher associated potentials.

The highest potential is selected as the first cluster center. Next, the potential of all the remaining points is reduced. Defining z_1^* as the first group center and denoting its potential as P_1^*, potential of the remaining points is reduced as in (10):

$$P_i \leftarrow P_i - P_1^* e^{-\beta\|z_i - z_1^*\|^2}, \ \beta = \frac{4}{r_b^2} \quad (10)$$

Where, the constant $r_b > 0$ defines the neighborhood radius with sensitive reductions in its potential.

In this way, points close to the selected center will have their potentials reduced in a more significant manner, and so the probability of being selected as centers diminishes. This procedure has the advantage of avoiding the concentration of identical clusters in denser zones. Therefore, r_b is selected in order to be slightly higher than r_a, so as to avoid closely spaced clusters. Typically, r_b=1.5r_a.

After performing the reduction of potential for all of the candidates, the one with the highest potential is selected as the second cluster. Then, the potential of the remaining points is again reduced. Generically, after determining the *r*th group, the potential is reduced as follows (11):

$$P_i \leftarrow P_i - P_r^* e^{-\beta\|z_i - z_r^*\|^2} \quad (11)$$

The procedure of center selection and potential reduction is repeated until reaching a good tradeoff between having a sufficiently high potential and being distant enough from the clusters determined before [13].

IV. Data Acquisition and Feature Selection

Clinical data about newborns were collected by filling approved blank sheets from delivery occurrences during three months in 2006 from two general hospitals in Tehran. Data acquired from clinical data, physicians' prescriptions and questionnaire from the newborn's family. This data includes both newborns affected by jaundice and the healthy ones. We removed some incomplete records. Of course because of different features for prediction of jaundice, number of removals was different for each model. Finally 552 records were collected, of which 515 and 333 records were used for prediction before and after delivery, respectively. There is a wide range of factors that affect neonatal bilirubin levels [14]. Selected features are as follows.

Features used for prediction before delivery include mother's O blood group (being O blood group or not), mother's having diabetes, mother's having hypertension, fetus' gender (males have more risk of jaundice), being the first child of family, siblings' jaundice background existence and mother's age. It is assumed that fetus's gender has been declared by ultrasonography.

Features used for prediction just after delivery include ABO incompatibility, Rh incompatibility, cesarean delivery, mother's regional anesthesia for delivery, mother's using of oxytocin, mother's hypertension, mother's diabetes, infant's gender, being the first child, prematurity, siblings' jaundice background, mother's age, infant's weight, gestational age.

V. Results

The clustering methods for prediction before and after delivery designed using fuzzy clustering toolbox of MATLAB software. Three clustering methods (SC, FCM and SFAM) used for this purpose. To choose the best architecture of each clustering method, it was trained and tested for different parameters, e.g. cluster radius for SC method, number of clusters for FCM method and vigilance value for fuzzy ARTMAP method. One of the most common evaluators to evaluate the performance of a classifier is called evaluation matrix. Four combinations of classifier output and desired output are possible for jaundice prediction, which is shown in Table III. Based on these parameters some criteria are calculated which follow. In these formulae TP, FP, FN, and TN are number of occurrence of corresponding state [15].

- *Sensitivity (SE):* the probability that a test result is positive given the subject has the disease. In a suitable experiment the sensitivity can be estimated by: TP/(TP+FN)
- *Specificity (SP):* the probability that a test result is negative given a subject does not have the disease. In a suitable experiment the specificity can be estimated by: TN/(TN+FP).

TABLE III
EVALUATION MATRIX

		Jaundice Presence	
		YES	NO
Classifier Output	YES	True Positive (TP)	False Positive (FP)
	NO	False Negative (FN)	True Negative (TN)

• *Correct Classification (CC):* the probability that the test result reflects the true disease state. In a suitable experiment the probability of a correct classifier output is estimated as the proportion of cases for which the classifier output is correct: (TP+TN)/(TP+FP+TN+FN).

The performance of a classifier should, wherever possible, be expressed in terms of sensitivity, specificity and Correct Classification. Table IV shows the results of jaundice prediction before and after delivery.

TABLE IV
RESULTS OF JAUNDICE PREDICTION

Neonatal Jaundice Prediction	Clustering Method	Time (sec)	Train CC%	Train SE%	Train SP%	Test CC%	Test SE%	Test SP%
Before delivery	SC	0.12	75.38	97.12	53.56	76.03	96.52	55.73
	FCM	4.42	64.05	66.14	59.07	70.97	73.64	64.36
	ARTMAP	65.04	80.07	78.69	83.37	64.44	80.98	53.45
After delivery	SC	0.98	100	100	100	81.27	88.36	66.57
	FCM	6.18	98.59	99.64	95.73	78.42	85.10	65.95
	ARTMAP	0.9	100	100	100	75.03	80.63	67.84

VI. DISCUSSION

High percentage of jaundice (70.83) may be due to Iranian race, the period of gathering data and the fact that almost all Iranian infants are breast-fed. The different number of data entries used in each model is due to different number of missing data for each feature. Since number of features for prediction before and after delivery was 7 and 15, respectively the total missing data is more in the latter. As can be seen results of prediction before delivery are not very satisfactory. This is due to insufficient information for this prediction. However considering this fact that there is not a clear method about jaundice prediction by physicians, even this level of prediction which is gained by fuzzy clustering is at least a good step toward solving jaundice prediction problem. In case of prediction after delivery these results are more acceptable because some additional risk factors about the infants are available.

VII. CONCLUSION

This research was done after collection of a total of 552 medical records from infants born during April to June 2006 in two general hospitals in Tehran. Because of some missing values for each input variable, impossibility of complete follow up and communication problems we had to remove some of incomplete record files. Of course because of different aspects for prediction of jaundice, number of removals was different for each model. Finally 515 and 333 files used for jaundice prediction before and after delivery, respectively. In this work we tried to predict neonatal jaundice before and after delivery. For this purpose three clustering methods (SC, FCM and fuzzy ARTMAP) are applied. The number of parameters in these models is different due to different available input parameters for each model. To evaluate results of these models we used evaluation performance matrix criteria, which include percentage of correct classification, sensitivity, and specifity (SP%). The above mentioned criteria for jaundice prediction before delivery were approximately 76%, 97%, 56%, while for jaundice prediction after delivery were 81%, 88%, 67%, respectively. These results show that the proposed clustering methods can achieve satisfying results for predicting risk of jaundice considering this fact that physicians do not have any estimation about probability of jaundice appearance.

ACKNOWLEDGMENT

Authors acknowledge their best for guides and cooperation from Mrs. Vatanparast, Dr Talebian Moghaddam, Dr Ghadiani, Mrs Nayshabouri for helping in data gathering, Management & Staff of Baqiyatallah & Najmieh Hospitals and everybody who helped or encouraged the authors.

REFERENCES

[1] Vinod K. Bhutani, Glenn R. Gourley , Saul A, Bili K, "Noninvasive measurement of total serum bilirubin in a multiracial predischarge newborn population to assess the risk of severe hyperbilirubinemia", Pediatrics,Vol.106, No.2, August 2000

[2] Beeby, Philip (2003). Jaundice, RPA Newborn care protocol book, Royal prince alfred hospital

[3] Charles Lee, Tram Hoang, Shazia Khan, "Neonatal jaundice-its mathematical model and treatments", Department of Mathematics, California State University Fullerton.

[4] Andreas Seyfang, Silvia Miksch, Mar Marcos, "Combining diagnosis and treatment using Asbru".

[5] American Academy of Pediatrics, "Management of hyper-bilirubinemia in the healthy term newborn", Vol.94, No. 34, 1994.

[6] M.Bottini, G.F. Meloni, F.Gloria Bottini, "Seasonal pattern of phototherapy: a study in the Sardinian population", Biological Rhythm Research, Vol.34, No.1, pp 13-21, 2003.

[7] David K.Stevenson, Avroy A. Fanaroff, Jeffrey Maisels, et al, "Prediction of hyperbilirubinemia in near-term and term infants", Pediatrics, Vol.108, No.1, July 2001.

[8] R.Khayati, F.Towhidkhah, F.H.E Tehrani, "Modeling of jaundice treatment using stateflow", 10th Iranian Conference on Biomedical Engineering (ICBME 2004), p 50, In Persian.

[9] A.Ahouraei, F.Towhidkhah, R.Khayati, "Diagnosis of neonatal jaundice using ANN", 11th Iranian Conference on Biomedical Engineering (ICBME 2005), p 70, In Persian.

[10] George E. Tsekouras, Haralambos Sarimveis, "New approach for measuring the validity of the fuzzy c-means algorithm", Advances in Engineering Software, Vol. 35, 2004, pp 567–575.

[11] Jervis, B.W., Garcia, T. (1999). Probabilistic simplified fuzzy ARTMAP, *IEE Proc.-Sci. Meas. Technol., Vol. 146, No. 4*, 165–169.

[12] S.L. Chiu, "Fuzzy model identification based on cluster estimation", J. Intell. Fuzzy Systems 2 (1994) 267–278.

[13] Rui Pedro Paiva, Antonio Dourado, "Interpretability and learning in neuro-fuzzy systems", Fuzzy Sets and Systems, Vol. 147, 2004, p 17.

[14] M. Jeffrey Maisels, "*Jaundice*, neonatology, pathology and management of the newborn", by G. B. Avery, Lippincott William & Wilkins, Ch. 38, pp. 765-770

[15] Tianxi Cai Margaret, S. Pepe Thomas Lumley, Yingye Zheng, "The sensitivity and specificity of markers for event times", UW Biostatistics Working Paper Series, University of Washington, Paper 188, 2003.